InTEXT Systems: Unique, Intelligent Tools for the Internet

Only *InTEXT*'s powerful Software Development Kits (SDKs) provide the next generation value-add that today's users expect from Internet and Web products. *InTEXT*'s unique technology provides the time-to-market advantage you need to help knowledge workers and consumers benefit most from your product:

- *Profiling with natural english*—easily create user profiles with simple phrases and seed sentences
- *Self-tuning agents*—employ learning agents to monitor new information as it becomes available
- *Summarizing*—create document abstracts *on the fly*: no indexes, no topic trees, no semantic networks
- *Automatic HTML & hyperlink generation*—quickly & easily publish documents to the Web
- *HTML-to-HTML search & retrieval*—search HTML documents & retrieve them with all links intact
- *Natural language querying*—find crucial information: no complicated query structures to learn
- *Relevancy profiling & retrieval based on content*—know what information most applies, instantly
- *Multilingual engines*—French, German, Spanish, Chinese and more...
- *Multi-platform support*—Windows & major UNIX platforms
- *Scalable architecture*—SDKs deployable with browsers, servers & on-line services

InTEXT Systems' Intelligent Internet Development Tools are based on a powerful Heuristic/Learning architecture, which uses skim-reading, comprehension and self-tuning techniques to understand the content of information. Where other technologies force developers to perform extensive indexing and create tremendous overhead, *InTEXT*'s technology provides real-time performance with little or no indexing and minimal overhead:

InTEXT Object Router Agent SDK

Automatically monitors live data, including news groups and wire feeds, and channels it directly to users' desktops without requiring indexing.

- *On-the-fly information profiling*
- *Natural language filter creation*
- *Dynamic, self-tuning filters*
- *Scalable, client/server toolkit*

InTEXT Object Analyzer Agent SDK

Creates document abstracts on the fly, reducing network overhead and saving users from information overload.

- *On-the-fly document summarizer*
- *Automatic keyword generator*
- *Internal hypertext linking*
- *Powerful application development base*

InTEXT Precision Agent SDK

Reduces full-text indexes tenfold, saving disk space and increasing both speed & accuracy significantly.

- *Dramatically reduced server indexes*
- *Rapid retrieval and unparalleled accuracy*
- *Powerful platform for application development*

InTEXT Retrieval Engine, WebServer SDKs

Delivers instant access to the world of on-line information by supporting scalability, extensibility and compatibility.

- *Powerful search, retrieval and storage APIs*
  - *WAIS, Z39.50 and HTML compliance*
  - *Document mgmt, RDBMS & 4GL extensibility*
- *Dynamic relevancy ranking*
- *Several search facility options*
  - *Full boolean, fuzzy, phrase, proximity, free form english searches and many others...*
- *Native document display and output, incl. HTML*
- *Multiple index and security levels*
- *Multilingual, incl. European & Asian languages*
- *Windows, DOS, UNIX and Mainframe support*
- *Client/Server architecture supporting TCP/IP and SNA LU6.2 connectivity*

InTEXT NLQ SDK

Lets all users access the information they need using plain english and seed sentence queries.

- *Minimal user learning curve*
  - *Access information with "normal" language*
- *Promote on-line searches*
- *Intuitive for content-based searches*
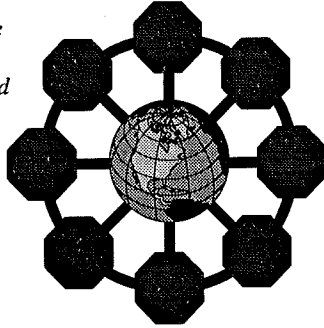- *Integrates with 3rd party search engines*

Call today for a free white paper: *Tools for the Internet: Channeling the Information Tidal Wave*
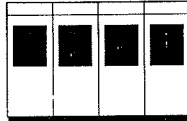
CD, Diskette,
Facsimile,
OCR, On-Line
News Feeds,
Internet, World
Wide Web

Document,
Record,
Manual,
Research
Material

Mainframe,
Disk

Word Processor,
Terminal,
Printer

THE *InTEXT* RETRIEVAL ENGINE IS A CONTENT-BASED DOCUMENT SEARCH AND RETRIEVAL ENGINE FOR ORGANIZATIONAL TEXTBASES, THE INTERNET AND THE WORLD WIDE WEB. IT SUPPORTS UNIX, MVS/CICS, MVS/TSO, VAX/VMS, OPEN VMS, WINDOWS AND MS/DOS.

"The InTEXT Retrieval Engine is the most powerful and diverse document storage and retrieval product available in the industry today."

Internet and Web Search Engine

The *InTEXT* Retrieval Engine is WAIS (Wide Area Information Server), Z39.50 and HTML compliant and can be used as a search engine for the Internet and the Web. For example, developers can use *InTEXT* to provide powerful query capabilities to Internet and Web textbases so users can find relevant on-line information.

Dynamic Document Search and Retrieval

Free-form natural language querying, full Boolean searching and positional relationship, the *InTEXT* Retrieval Engine has it all. From Windows, UNIX clients or character terminals, users can formulate both structured and free-form English queries to initiate on-line information searches.

Dynamic Relevancy Ranking

The *InTEXT* Retrieval Engine provides the most sophisticated retrieval and document ranking mechanism available today. Based upon a content understanding architecture, the Retrieval

InTEXT

Engine automatically self-tunes to the document collection as it changes.

It reflects alterations in the relevancy percentages assigned to documents, which are presented to users in new priority orders.

Further, users can assign individual weights to their search criteria to subjectively influence document priority ranking.

Native Document Display and Output

Once users find what they want, the *InTEXT* Retrieval Engine provides many ways to view the results. Users can look at a list of document titles, sort them, natively view those pages that contain a search term or output the results to a file or printer.

Enterprise-Wide Solution

If your organization is lost in a document-intensive world, the *InTEXT* Retrieval Engine is a truly cost-efficient solution. Running across the enterprise, *InTEXT* turns the mass of electronic information into a business advantage.

Search Facilities:
- *Full Boolean Logic*
- *Free Form Natural Language*
- *Structured Data (e.g. dates, names)*
- *Fuzzy Logic*
- *Phrase*
- *Proximity*
- *Stemming and Wildcard*
- *Synonyms*
- *Structured Thesaurus*
- *User Weighted Search Terms*
- *Unlimited Search Scoping*

Display Facilities:
- *Automatic Results List Ranking*
- *Search Term Highlighting*
- *Document Display by Title, Search Term in Context, Smart Full-Text, Document Parts*
- *Output Results to Terminal, File, Printer, Other Applications*

System Features and Facilities:
- *Powerful Macro Language*
- *User Customizable Search Screens*
- *Inverted Index, Built-in Compression*
- *Dynamic Real-Time Indexing*
- *Strong Security*
- *Batch Processing*
- *On-Line Help*
- *Flexible Textbase Structure*
- *Client/Server Architecture*
- *Full-Featured API*

Features	Benefits
Natural Language Query	• Access info with "normal" language • No new languages to learn • Minimal employee training • Eliminate manual searches
Dynamic Index Management	• Easy-to-change indices • Dynamic textbase update • Tremendous timesaver
Definable Index Levels	• Objective specification • Subjective specification • Customization • Ease of use
Powerful Application Development Base	• Flexible development with macros • "Learning" applications • API Library for customization • Dynamic rather than static • Easy to change, upgrade, administer • Client/Server architecture provides enterprise-wide scalability • Supports popular textbases such as STATUS™
World Wide Web and Internet Search and Retrieval	• Supports WAIS, Z39.50 and HTML protocols

---

InTEXT

Author Station
ADD, REVISE, DELETE, FIX, MIGRATE

LAN

Retrieval Station
SEARCH, DISPLAY, PRINT

Enterprise Document Repository

CICS Enterprise Manager

DB2 Control Table

InTEXT WebServer | 3270 Display

Local Document Repository

OS/2 Enterprise Manager

Database Manager Control Table
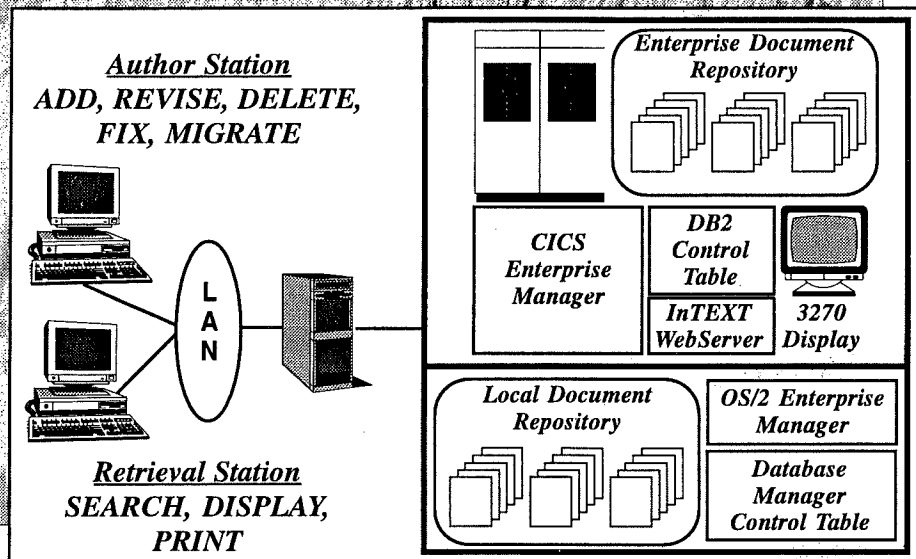
THE *InTEXT* ENTERPRISE MANAGER PROVIDES DOCUMENT STORAGE, MANAGEMENT, AND RETRIEVAL ACROSS AN ENTERPRISE ENVIRONMENT INCLUDING MAIN-FRAME, WINDOWS, OS/2, UNIX AND CHARACTER WORKSTATIONS.

"InTEXT Enterprise Manager provides the only desktop-LAN-mainframe based distributed document manage-ment solution available today."

InTEXT

Enterprise-Distributed Applications

Building Windows-based applications that communicate with IBM MVS/CICS has never been easier. *InTEXT* Enterprise Manager is the solution for creating distributed document management applications. With *InTEXT* Enterprise Manager, users have full access to Document Management functions such as full check-in/check-out management and document migration across LANs and WANs—all from their desktops.

Document Management

Besides mainframe-to-desktop communication, Enterprise Manager has a powerful set of application programming interfaces (APIs) for managing large volumes of documents and provides the ability to add, update, delete, get, fix and migrate documents. Documents can reside on the LAN and are known to all LANs in the WAN. They can also reside on the mainframe and are available to all servers.

Full-Text Retrieval

Enterprise Manager is compatible with many popular text retrieval prod-ucts, including STATUS™. Documents that are moved to the mainframe

can be added to or deleted from STATUS textbases. These documents are then available for full-text retrieval using STATUS either via character based terminals or GUIs.

Built-In Support for Communication Protocols

With Enterprise Manager, there are no more networking worries. All of the protocols are built in. Whether using SNA APPC or TCP/IP at the mainframe level or Token Ring, TCP/IP or Named Pipes at the server level, communication and document transfer occur seamlessly.

Portability

The *InTEXT* Enterprise Manager can be used in a number of configurations. Please discuss any special requirements with your *InTEXT* Systems representative.

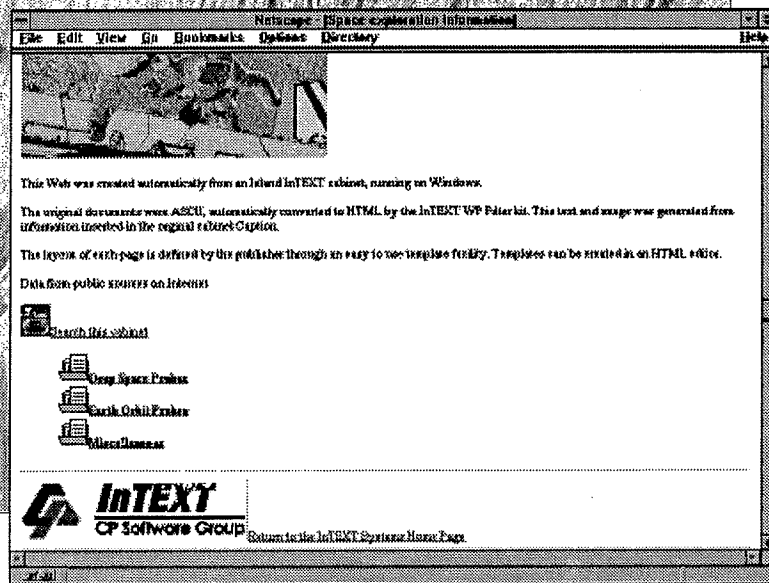| Features	Benefits
MVS to Windows Tools	• Distributed application development toolkit for enterprise-wide document management • Leverage existing investment in legacy applications • Access all mainframe information assets
Document Management	• Complete check-in/check-out, add, update, delete, get, fix, migrate functionality for full-scale document control • Enterprise-wide, heterogeneous environment for seamless sharing of documents
Open Architecture	• Can be ported to other platforms and other LAN environments • Supports multiple communication protocols • Compatible with *InTEXT* WebServer and STATUS for Internet and Web document publishing, search and retrieval

InTEXT

THE *In*TEXT WEBPAK IS THE WORLD'S FIRST SUITE OF PRODUCTS AND TOOLS FOR ADDING INTELLIGENCE TO WEB-BASED DOCUMENT PUBLISHING, PROFILING, HTML-TO-HTML SEARCH & RETRIEVAL, NATURAL LANGUAGE QUERYING AND SUMMARIZATION FOR BOTH THE UNIX AND WINDOWS ENVIRONMENTS.

"The WebPak Client/Server Development Environment helps organizations quickly create intelligent Web sites, publish Web pages at the click of a button and efficiently leverage their information assets in the on-line world."

InTEXT

Intelligent Web Site Deployment

Combining a Windows-based automated publishing environment, Web-ready applications for rapid deployment and a well-versed combination of Software Developer's Kits (SDKs), the *InTEXT* WebPak provides Web-site developers with intelligent tools for quickly deploying advanced Web-based applications.

Automated Web Publishing—Island *InTEXT* 2.0

Using the WebPak, application developers can build Web publishing sites that support popular Web browsers, such as Netscape and Mosaic, WAIS and Z39.50-compliant Internet clients and leading full-text retrieval databases.

These Web site applications can support several unique capabilities, including free-form English user queries, live document routing, real-time summarization, content-based searching, relevancy-ranked retrieval, automatic hyperlink and keyword generation and automatic Web creation.

InTEXT Object Router SDK

Included in the WebPak kit, the *InTEXT* Object Router Intelligent Agent SDK automatically monitors live data, including news groups and wire feeds, and channels it directly to users' desktops without requiring indexing.

InTEXT Object Analyzer SDK

This document summarizing toolkit creates document abstracts on the fly, reducing network overload and saving users from information overload.

InTEXT WebServer SDK

Based on the powerful *InTEXT* Retrieval Engine, the WebServer delivers instant access to the world of on-line information by supporting scalability, extensibility and compatibility. The WebServer provides relevancy-ranked retrieval, based on document content, and supports HTML, WAIS & Z39.50 protocols.

InTEXT NLQ SDK

The *InTEXT* NLQ Engine is the most advanced free-form English facility of its kind. It lets users access the information they need using plain english and seed sentence queries, minimizing and even eliminating syntax training.

Intuitive Architecture

The *InTEXT* WebPak products and SDKs are based on *InTEXT* Systems' Heuristic/Learning™ architecture that analyzes document surface structures to gain content knowledge of live and stored information. This unique architecture makes the *InTEXT* technology critical tools for the highly interactive and changing content of the Internet and the Web.

Features	Benefits
Automated Web Publishing	• HTML conversion layer and automatic hyperlink generation let *anyone* create Web pages
Profiling with Natural English	• Easily create user profiles with simple phrases and seed sentences
Self-Tuning Agents	• Employ learning agents to monitor new information as it becomes available
Summarizing	• Create document abstracts *on the fly* and avoid building indexes, topic trees and semantic networks
Relevancy Profiling and Retrieval	• Users instantly know what information most applies to their information needs
World Wide Web and Internet Search and Retrieval	• Supports WAIS, Z39.50 and HTML protocols
Scalable Architecture	• SDKs deployable with browsers, servers and on-line services

InTEXT

Automatic Message Router processing example.msg

File   Window   Option   Help

FILTER / USER	JONES	CAROL	ANDREA	ADRIAN	
DUKEQA	0				
WAR		0			
COMPANY		3			
ENGINES	1				
IRAQ		5		5	
STOCKS	5		16	16	
MONEY					
NEURAL	0			0	
FLOODS				4	
SOFT				1	
TAMPA					

View

Num	Title
052	THE BELL GROUP LTD - ANNOUNCEMEN
053	MIM HALF YEAR PROFIT SLUMPS, ANNU
054	FRANKFURT EXCH OPG ONE - JAN 16

User - JONES, Filter - STOCKS

View

Score	ID	Title
100	003	NIKKEI DOWN, DLR, _5 TOKYO
100	012	AUST STOCKS CLOSE WEAKER
100	045	DORAL RESOURCES NL - Notic
100	050	Auspharm International Limited-
80	013	TOKYO STOCKS _5 TOKYO

THE *InTEXT* OBJECT ROUTER INTELLIGENT AGENT TOOLKIT ROUTES LIVE AND STORED INFORMATION TO USERS BASED ON THEIR INTEREST PROFILES—WITHOUT REQUIRING INDEXING. OBJECT ROUTER IS A POWERFUL TOOL FOR MANAGING THE TIDAL WAVE OF ON-LINE INFORMATION.

"Object Router's content-based approach streamlines the organization's business process by using intelligent agents to route information based on a user's definition of what's important."

-Delphi Consulting, October 1994

InTEXT

On-the-Fly Information Routing

With the Object Router technology, users can engage intelligent agents to search their organization's databases and on-line news feeds for specific data and dynamically route it to user-defined folders for easy access and usage.

Object Router allows users to identify, in natural language, the topics in which they are interested. As an intelligent database and network agent toolkit, Object Router simply watches for those topics and brings them to user desktops.

Filter Intelligence

Object Router filters learn from the stream of documents and dynamically tune themselves to better discriminate between relevant and irrelevant messages each time they are confronted with new subjects.

The filters become more effective as they mature and become more sensitive to rapidly changing issues, giving users faster access to the most critical information, automatically.

Before being routed to users' electronic mailboxes, documents are ranked based on the material's relevancy to user-defined subjects. Users have the flexibility to specify how much priority a specific subject should have compared to other filter subjects.

A single filter can serve multiple recipients, each with a different relevance threshold. Some users may wish to see messages if they are central to the topics covered in a filter, while others may adopt a more general information gathering approach. Object Router caters to the various needs within workgroups.

On-Line Profiling and Routing Solutions

Object Router is part of *InTEXT* Systems' complete line of content management solutions. It can be used by corporations and institutions to create customized information profiling and management solutions for their own databases and for the Internet.

Open Architecture

The Object Router toolkit is a C language subroutine library that allows programmatic access to Object Router's key features. It is available for UNIX and Windows.
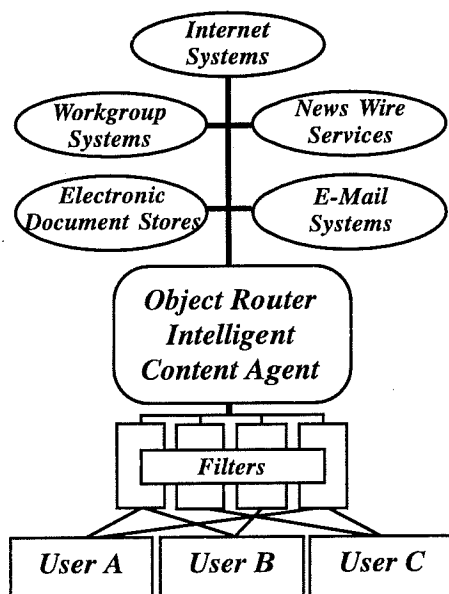
The Object Router Intelligent Content Agent is an electronic information sleuth that scans organizational and Internet data and routes it to appropriate users.

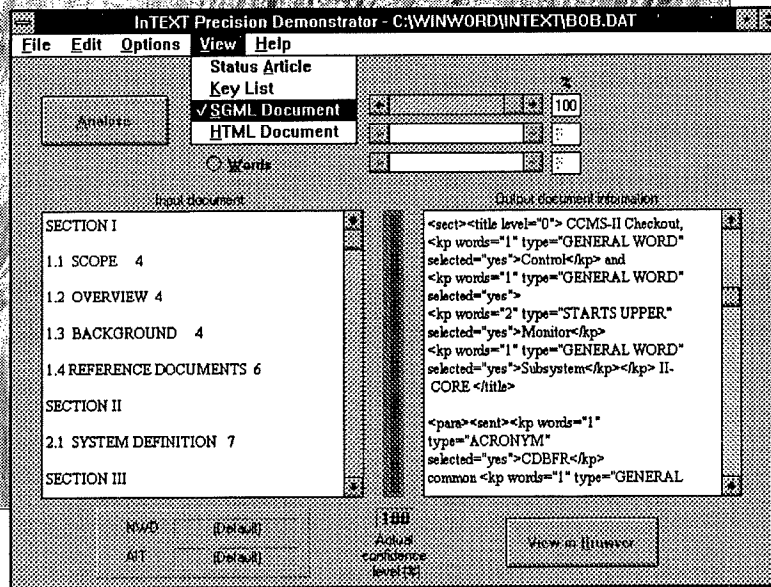| Features	Benefits
On-the-Fly Routing	• No indexing required • The most appropriate documents automatically route to interested users • Incoming e-mail automatically sorted into topic folders
Filtering/Routing/Classification	• More timely access to current information • Access to only the documents people want • Reduces information overload • Reduces network traffic
Filter Intelligence	• Filters that learn from the stream of information and automatically tune to the changing nature of information • Users receive leading edge information

InTEXT

I_nTEXT_ PRECISION IS AN ACCELERATED INDEXING AND DYNAMIC CONTENT AGENT TOOLKIT USED FOR SIGNIFICANTLY REDUCING FULL-TEXT INDEXES AND AUTOMATING HTML & SGML DOCUMENT PREPARATION FOR ON-LINE PUBLISHING.

"Precision's automation of the document preparation process for on-line publishing and retrieval vendors is extremely financially and commercially viable. What were formerly labor intensive activities, such as keyword generation and HTML conversion, are automatic with Precision."

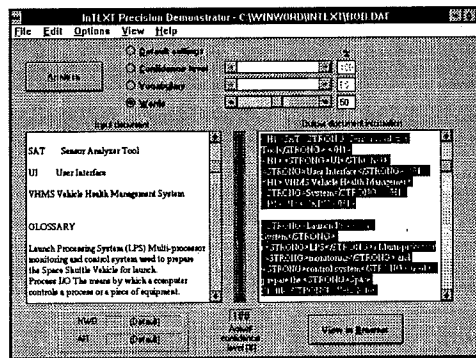## Intelligent Keyword Generation Facility

InTEXT Precision generates high-quality keywords from documents without bias. For example, in an authoring system, an application can use Precision to automatically fill in the keywords rather than forcing the user to do so. Traditionally, this valuable text retrieval technique involves a labor-intensive process, often frustrating end users and resulting in inaccurate search criteria. Precision's automation of this process makes it much more user friendly and commercially viable.

Index Reduction for High Precision and Recall

InTEXT Precision utilizes a unique PreciseScoping™ technology that generates a significantly smaller full-text index than any other commercial product. Precision does this by automatically creating content-aware, summarized documents.

These unique, summarized documents can be used in any high precision full-text retrieval application, enabling a significant increase in precision while producing indexes that are 5 to 10 times smaller.

InTEXT

InTEXT Precision is unique in the way it tackles precision retrieval and recall from the indexing end rather than the querying end. The summarized documents, with the most content-bearing words intact, are indexed rather than the entire document. Because of the dramatically smaller index, retrieval speed is improved. The content knowledge of the summarized document improves relevancy ranking or "hits" of documents that fit a particular query.

InTEXT Precision offers all full-text retrieval vendors and on-line publishers the opportunity to dramatically improve the accuracy of their retrieval system without sacrificing recall.

Automatic Conversion to Standard Formats

While automatically converting documents into standard interchange and publishing formats, *InTEXT* Precision streamlines the creation of logical structure identifications and SGML, HTML and keyword tags. Additionally, Precision enables application developers to automate hypertext linking. This markup feature greatly reduces the efforts required in publishing to the World Wide Web and the Internet.

Software Development Kit

InTEXT Precision is available as an SDK, with C libraries and documentation, for UNIX and Windows operating systems.

The InTEXT Precision agent technology automatically creates content-aware, HTML marked up, and "precise indexed" documents.

Features	Benefits
Content-Aware Document Preparation	• Increased precision and recall in document search and retrieval • Documents ready for web clustering, routing, crawling to get users the most up-to-date information possible
Automatic SGML and HTML Conversion	• Documents instantly ready for Internet and Web publishing • Standardized for worldwide document exchange
Reduction in Index Size	• Storage overhead reduction • Faster search and retrieval • Maintenance reduction • Increase in computer responsiveness
Automatic keywording	• Reduces staff tasking significantly • Increases keywording accuracy • Future-proofs end user applications by automatically filling in word processing/ document authoring keyword prompts
Open Architecture	• Built on Heuristic/Learning architecture so content knowledge is inherent • API library for customization

InTEXT

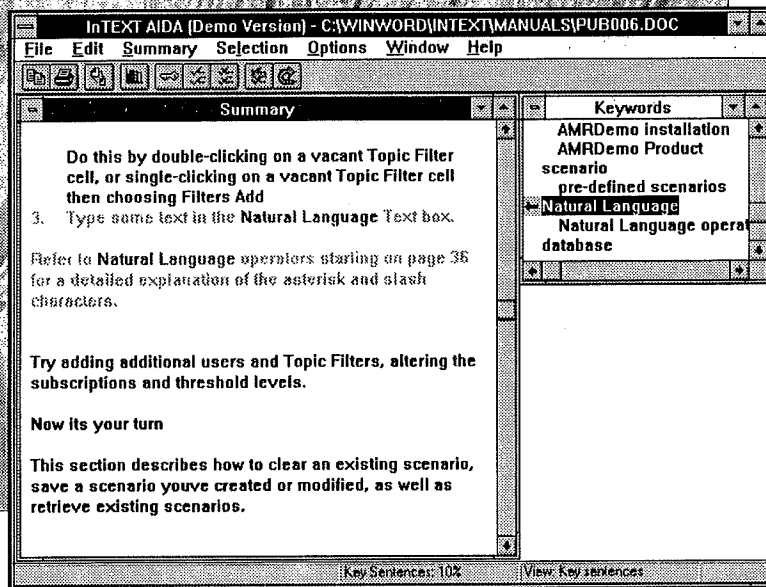**InTEXT AIDA [Demo Version] - C:\WINWORD\INTEXT\MANUALS\PUB006.DOC**

File   Edit   Summary   Selection   Options   Window   Help

Summary

Do this by double-clicking on a vacant Topic Filter cell, or single-clicking on a vacant Topic Filter cell then choosing Filters Add

3.   Type some text in the Natural Language Text box.

Refer to Natural Language operators starting on page 35 for a detailed explanation of the asterisk and slash characters.

Try adding additional users and Topic Filters, altering the subscriptions and threshold levels.

Now its your turn

This section describes how to clear an existing scenario, save a scenario youve created or modified, as well as retrieve existing scenarios.

Keywords
- AMRDemo Installation
- AMRDemo Product scenario
  - pre-defined scenarios
  - Natural Language
  - Natural Language opera
  - database

Key Sentences: 10%   View: Key sentences

InTEXT OBJECT ANALYZER PROVIDES REAL-TIME CONTENT ANALYSIS AND SUMMARIZATION OF DOCUMENTS. OBJECT ANALYZER IS A POWERFUL TOOL FOR ACCESSING THE MOST CRUCIAL PARTS OF INFORMATION, ACROSS THE ENTERPRISE AND THE INTERNET.

"InTEXT's useful summarizing tool gives users an excellent way to cut down on the sheer volume of documents. ...We found the heuristic algorithms InTEXT uses to create summaries to be quite sophisticated."

- PC Week, Oct '94

InTEXT

On-the-Fly Summarization

Object Analyzer identifies the major information content of large documents and creates precise synopses of those documents. Using *InTEXT's* unique heuristic technology, Object Analyzer understands the "aboutness" of documents and deduces the most important information on the fly. No front-loading of dictionaries or index building is required. Object Analyzer automatically creates summaries from 1-99 percent of the original.

Document Assimilation

Object Analyzer enables users to comprehend complex and lengthy documents by extracting the document's main focus. It performs analyses by reading the entire document and objectively rating every word, phrase and sentence by its contribution to the document's central theme or themes.

Dynamic Keyword Indexing

Object Analyzer automatically creates a coherent index of a document's key words and phrases. Users can display, save and

print this index, complete with word variation, word frequency and alphabetical listing options.

Content Knowledge Expansion

While Object Analyzer's content knowledge is built-in and domain independent, it can also be tuned to make more accurate and comprehensive extractions from documents associated with particular topics and disciplines. For example, medical

terms can be added and dynamically called upon during document analysis.

Automatic Hypertext Preparation

Object Analyzer generates automatic hypertext links from the keyword index to associated words within a summary. And, Object Analyzer can output abstracts in SGML format. It marks all instances of key phrases and key

sentences, and it represents key phrases grouped by their syntactic variants and sub-phrase families.

Open Architecture

For application customization, the Object Analyzer Software Development Kit (SDK) contains a C language subroutine library that allows programmatic access to the Object Analyzer's key features. The SDK is available for UNIX and Windows.
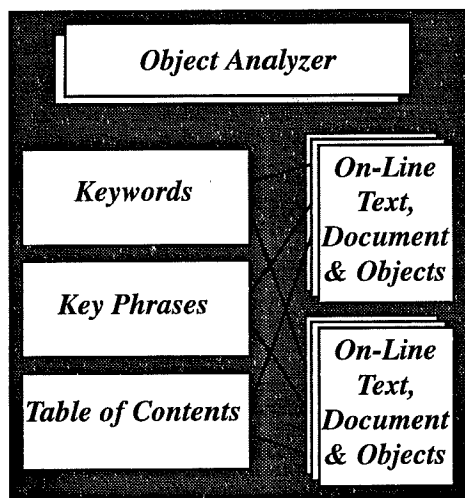
The Object Analyzer summarizing agent understands the content of documents and creates accurate summaries on the fly. It also creates keyword indexes automatically.

Features	Benefits
On-the-Fly Summarization	• Objective interpretation of documents • No time-consuming preparation • No dictionaries or indexing required • Instant access to the most important pieces of information • 1-99% summaries of the original • Read thumbnail versions of Internet and Web documents
Automatic Keyword Generation	• Easy keyword lookup for fast research • Documents already prepared for keyword assignments in text retrieval engine—tremendous timesaver • Internal hypertext linking and SGML generation for standardization ease • Can perform unique "next" and "last" keyword searches
Powerful Application Development Base	• Future-proofed enterprise investment • Efficient Internet publishing preparation and utilization • API library for customization